



SMARTA: Automated testing apparatus for visual discrimination tasks

Raymond Vagell¹ · Vance J. Vagell² · Rachel L. Jacobs³ · James Gordon¹ · Andrea L. Baden^{4,5,6}

Published online: 5 September 2019
© Psychonomic Society, Inc. 2018

Abstract

This article introduces the open-source Subject-Mediated Automatic Remote Testing Apparatus (SMARTA) for visual discrimination tasks, which aims to streamline and ease data collection, eliminate or reduce observer error, increase interobserver agreement, and automate data entry without the need for an internet connection. SMARTA is inexpensive and easy to build, and it can be modified to accommodate a variety of experimental designs. Here we describe the utility and functionality of SMARTA in a captive setting. We present the results from a case study of color vision in ruffed lemurs (*Varecia* spp.) at the Duke Lemur Center in Durham, North Carolina, in which we demonstrate SMARTA's utility for two-choice color discrimination tasks, as well as its ability to streamline and standardize data collection. We also include detailed instructions for constructing and implementing the fully integrated SMARTA touchscreen system.

Keywords Android application · Animal behavior · Data collection · Observer error · Experimental design · Ruffed lemurs · Color vision

The use of operant conditioning paradigms in animal research has allowed important insights into species' behaviors, cognitive abilities, and perceptual processing (e.g., McSweeney et al., 2014). The method, whereby an animal's behaviors are modified through reinforcement, was introduced by B. F. Skinner in the 1930s and is commonly conducted in an

operant conditioning chamber that bears his name (i.e., a “Skinner box”; Skinner, 1938). Although originally developed for research on rats and pigeons, over time the classic “Skinner box” has been modified by researchers for use in various discrimination tasks involving a wide range of species (e.g., pigeon: Emmerton & Delius, 1980; bat: Flick, Spencer, & van der Zwan, 2011; tortoise: Mueller-Paul et al., 2014; and dog: Müller, Schmitt, Barber, & Huber, 2015).

More recently, as technology continues to progress, there has been a push to develop more automated operant conditioning protocols using digital applications that are ideally also freely available. Some of the primary motivations behind these developments include (1) reducing costs of research, since a commercial “Skinner box” can cost in the thousands of US dollars (Chen & Li, 2017; Devarakonda, Nguyen, & Kravitz, 2016; Hoffman, Song, & Tuttle, 2007; Oh & Fitch, 2017; Ribeiro, Neto, Morya, Brasil, & de Araújo, 2018); (2) improving the speed and/or accuracy of data collection (Chen & Li, 2017; Oh & Fitch, 2017); and (3) taking into account the needs of the study and study subjects. Regarding the latter point, novel techniques have been developed to accommodate different testing environments (Fagot & Paleressompoulle, 2009; Oh & Fitch, 2017), discrimination tasks (Ribeiro et al., 2018), and rewards (Steurer, Aust, & Huber, 2012).

Thus, a number of operant conditioning systems have recently been developed that take advantage of technologies, such as touch screens and microcontrollers, to provide options

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-018-1113-9>) contains supplementary material, which is available to authorized users.

✉ Raymond Vagell
raymond.vagell@gmail.com

- ¹ Animal Behavior and Conservation Program, Department of Psychology, Hunter College of the City University of New York, New York, NY, USA
- ² Queens, USA
- ³ Center for the Advanced Study of Human Paleobiology, Department of Anthropology, George Washington University, Washington, DC, USA
- ⁴ Department of Anthropology, Hunter College of the City University of New York, New York, NY, USA
- ⁵ Departments of Anthropology and Biology, The Graduate Center of the City University of New York, New York, NY, USA
- ⁶ The New York Consortium in Evolutionary Primatology, New York, NY, USA

that require low financial investment and/or technological expertise on the part of researchers. These systems, such as ArduiPod Box (Pineño, 2014), ROBucket (Devarakonda et al., 2016), CATOS (Oh & Fitch, 2017), and OBAT (Ribeiro et al., 2018), are designed to detect different behavioral actions, such as a nose-poke (Devarakonda et al., 2016; Pineño, 2014), button press (Oh & Fitch, 2017), or bar touch (Ribeiro et al., 2018), in response to visual (Pineño, 2014) or auditory (Oh & Fitch, 2017; Ribeiro et al., 2018) stimuli.

Unfortunately, many of these systems were developed for small-bodied mammals (e.g., rodents, callitrichine primates; Chen & Li, 2017; Devarakonda et al., 2016; Ribeiro et al., 2018) and may not be appropriate for larger-bodied animals. Moreover, there are additional drawbacks to some of these applications. For example, some systems utilize components that can require more specialized programming knowledge, such as PIC microcontrollers (Hoffman et al., 2007), or use hardware that is no longer available, such as the iPod touch (Pineño, 2014). Others use CSV files for data logging (Devarakonda et al., 2016; Oh & Fitch, 2017; Ribeiro et al., 2018), which can be cumbersome and necessitate constant data file downloads. Some systems are also designed to provide particular food rewards that might not be appropriate for some taxa (e.g., liquid; Devarakonda et al., 2016; Ribeiro et al., 2018). Although there may be no “one-size-fits-all” system for experimental animal behavior research, the continued development of open-source and automated systems is important for providing options to researchers, especially for those working with animals that have been less commonly used in such research.

To this end, we designed the Subject-Mediated Automatic Remote Testing Apparatus (SMARTA). SMARTA is a fully integrated touchscreen operant apparatus that utilizes state of the art technology for visual discrimination tasks. SMARTA offers advanced features such as online data logging, Bluetooth remote control, integrated video capture, and an automated food dispensing system. SMARTA is inexpensive and easy to build. Required parts and circuitry can be obtained “off the shelf” (supplemental material, Appendix A). The software was specifically developed for color vision research in captive lemurs, but is open-source and could be modified to display any visual stimulus, such as numbers, shapes, or faces with additional effort by an Android programmer. Such modifications would primarily be in the visual display components of the software, whereas core capabilities, such as data logging, Bluetooth remote control, and communication with an Arduino microcontroller for food dispensing could be reused with minimal modification. The hardware assembly, designed for medium-sized (~4–5 kg) animals, would be particularly useful for studies conducted in larger-bodied species than those used in studies involving the more classic “Skinner box.” Here, we describe the utility and functionality of SMARTA, as well as detailed instructions on how to assemble it.

Requirements

Hardware

SMARTA is easily assembled from consumer parts, which include four primary components: (1) an Android tablet, (2) an Android smartphone, (3) a motorized food delivery system, and (4) the apparatus cabinet (Figs. 1 and 2). A complete list of parts is provided in the supplemental materials, Appendix A. SMARTA costs around \$600 if constructed from new materials (as of 2018), but it can be significantly cheaper if users already have access to an Android tablet and phone.

Android tablet The tablet touchscreen is the key component of SMARTA, as it displays the stimuli and reacts to touches from animal subjects. SMARTA is designed to use a 10-in. Android tablet, such as the LG G Pad 10.1; however, any Android tablet that supports removable memory cards could be used. Since SMARTA records video, memory card capabilities will facilitate saving, downloading, and using the many video files it will produce. The tablet must also support Bluetooth and USB host functionality. The USB host capability, also known as USB OTG (i.e., “on the go”), is important because it allows the tablet to communicate with the Arduino-brand microcontroller to dispense food. Note that the tablet application assumes a 10-in. tablet. To use a different tablet size, SMARTA software would need to be updated with new tablet dimensions for touch targets, as well as for the overall user interface (UI) layout. A tablet running Android’s Lollipop release (5.0) or higher is strongly recommended, because this release includes the ability to “pin” an application, making it difficult to accidentally exit the application if animal subjects touch system buttons, such as “home” or “back.” Earlier Android versions can be used, but some animals may accidentally exit the application through errant swiping and tapping.

When the SMARTA tablet application is started, it waits for a Bluetooth connection from the SMARTA phone application (i.e., the Bluetooth remote control, described below). Once the connection is established, the SMARTA tablet application waits for various predetermined commands from the phone application to begin. Normally, communication is one-way from the phone to the tablet, with the tablet simply waiting for commands and executing the proper display or task flow without sending data back. The important exception is at the termination of training and testing modes, when the tablet sends data back to the phone application.

The tablet does not connect to the internet, nor does it require an internet connection. Rather, it listens for Bluetooth messages from the phone application and occasionally sends testing data via Bluetooth. Accordingly, the tablet does not require cellular data service or a wi-fi signal in order to function during testing. Since the phone application likely already has cellular data service, all online data logging is performed via the phone application.

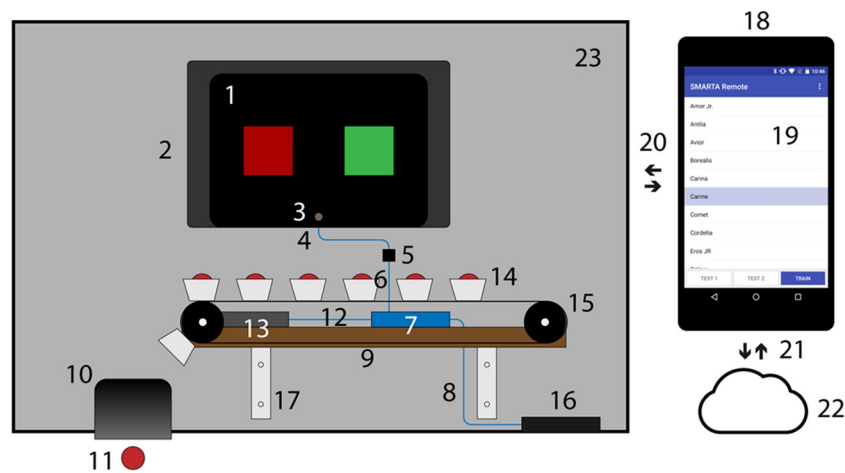


Fig. 1 Schematic illustration of SMARTA, with the front of the acrylic cabinet cut away to show the inner components, consisting of: (1) the SMARTA tablet app, (2) an Android tablet, (3) the integrated camera, (4) a USB OTG cable, (5) a serial-to-USB converter (FTDI), (6) a serial cable (jumper wires), (7) the microcontroller (Arduino with motor shield), (8) a USB power cable, (9) a wooden board, (10) an acrylic chute, (11) the food

reward (dried cranberry or raisin), (12) power wires, (13) a DC motor and gearbox, (14) a small plastic cup, (15) a conveyor belt, (16) a USB power pack, (17) a mounting bracket, (18) an Android phone, (19) the SMARTA phone app, (20) Bluetooth communication, (21) the Google Drive SDK (via the internet), (22) an online data spreadsheet, and (23) the acrylic enclosure (shown cut away).

Android smartphone SMARTA uses an Android smartphone, such as the Nexus 5 or Sony Xperia Z3, which acts as its remote control. The SMARTA phone application contains the bulk of the logic for how SMARTA operates. Generally speaking, both the Arduino microcontroller and tablet application merely execute commands that the phone application dictates. Furthermore, the phone application receives testing data from the tablet and is responsible for storing it online in Google Sheets. The application is organized into two main sections, training and testing, which mirror the two modes in which SMARTA may be used. The former mode is used to orient subjects to SMARTA and teach them to touch the object target, and the latter mode is used to collect data from subjects. The application also offers various configuration options, including an account picker (for phones with multiple user accounts), stimulus configuration mode, and the ability to choose among different Google Sheets spreadsheets for data logging. Setting values are stored in the phone's internal memory. As a result, different phones may use different accounts,

stimuli configurations, and spreadsheets with the same tablet and food dispenser, allowing SMARTA to be shared between researchers or projects.

The phone application has local data storage for when internet access is unavailable. When testing data are sent from the tablet to the phone via Bluetooth, the phone application first attempts to log the data online via Google Sheets and, if this fails, it stores data locally. The next time the application is started, if any local data are discovered, the user is prompted with the option to retry uploading the data online. In this way, data are never lost, and they will be uploaded online once an internet connection is available. Automated online data logging avoids the need to manually download CSV files. Because data can be seen in real time, erroneous entry can be identified more quickly. Moreover, this feature allows multiple trials to run on several SMARTA apparatuses while logging the data simultaneously to a single master spreadsheet. Local backup data may also be accessed via a text file on the phone, which is available via a USB connection.

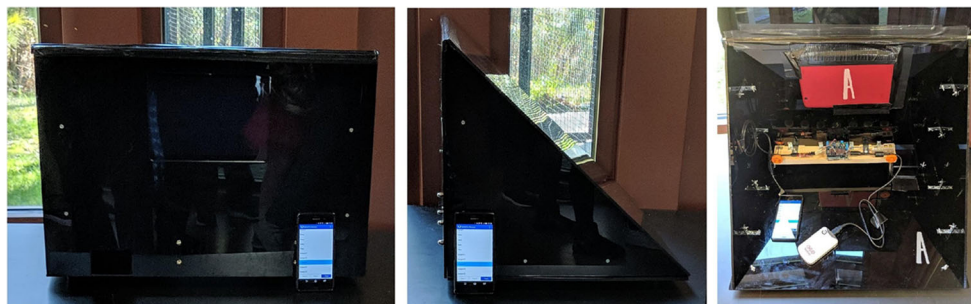


Fig. 2 SMARTA once fully constructed: front view (left), side view (middle), and inside view (right).

Motorized food delivery system SMARTA uses a motorized system to dispense food rewards to animals through a chute. SMARTA's food dispenser is built from a repurposed Tamiya-brand tracked vehicle kit (normally used to build toy tanks), which acts as a conveyor belt for food rewards. The conveyor belt is powered by a small motor that is controlled by an Arduino Uno microcontroller, fitted with a Adafruit Motor Shield version 2.3. The belt has seven small cups attached, each of which researchers load with a single food reward. These cups can be made from bottle caps or any other small dish, including those that can be 3-D printed or are commercially available.

Whenever the microcontroller receives a signal from the tablet, it advances the conveyor belt just enough so that a single food cup pours out its contents. The microcontroller receives signals from the tablet via serial communication. Because the tablet supports a USB host connection, a USB-to-serial converter can be used (e.g., SunFounder FTDI-brand Basic six-pin USB to TTL). Jumper wires are used to connect the microcontroller's general-purpose input/output (GPIO) pins to the FTDI-brand USB to serial converter, and a USB cable connects the FTDI-brand converter to the tablet. This enables standard serial communication between the tablet and microcontroller. Simple numerical values are sent from the tablet application to the microcontroller application to execute various conveyor movements. For example, the conveyor may be advanced one unit, or reversed one unit (to facilitate resetting the food cups for refilling) depending on which command is received via serial connection. The microcontroller is programmed with precise timing values for motor speed and the duration to turn it on, such that advancing the conveyor belt dispenses a single food reward.

The motor shield takes up all GPIO pins on the Arduino-brand microcontroller board by default, because it was designed to control a number of motors at once. Since SMARTA only uses one motor, many GPIO pins are actually still available. In a default SMARTA configuration, the microcontroller communicates via GPIO Pins 9 and 5 for transmit (TX) and receive (RX), respectively, which are passed through the motor shield. These pins are needed for the tablet's Android app to send serial commands to the microcontroller.

SMARTA's motor uses so little power that it can be powered via the Arduino-brand microcontroller board's own USB port. Normally, motor shields require higher current (e.g., provided via a separate power source) to ensure that the microcontroller has sufficient current even while the motor is activated. SMARTA uses a single motor at a slow speed, and powering it via USB does not cause any resets or other errors. As a result, any USB power source can be used that does not have an auto-off feature. Because SMARTA requires no external power source, one benefit of this system is that it is portable and can be used even in situations in which electrical outlets are unavailable.

The microcontroller and food conveyor are mounted on a wooden board, to give tension to the conveyor belt and keep all the moving parts and electronics organized within the acrylic cabinet. The board is attached to the cabinet with mounting brackets to carefully align the food reward cups with the exit chute for dropping food to animal subjects (Fig. 2).

This system was developed in collaboration with animal trainers at the Duke Lemur Center and was centered on the needs of our particular study subjects, which are highly food-motivated frugivores. Specifically, we developed a cup-based conveyor system to dispense sweet sticky rewards, such as craisins, which are preferred foods of the lemurs. However, a significant trade-off is the maximum capacity of seven food rewards per session. Other rewards, such as wheat grains or primate chow, were unsatisfactory for this study's subjects, but any reward that can be dispensed by a cup could be used. If a different food dispensing mechanism is preferred, SMARTA's Arduino software would need to be reprogrammed, and additional hardware might be required. For example, adding a food hopper or juice dispenser would not use the Arduino motor shield, and would instead require additional custom electronics specific to activating it (e.g., additional power supply and relay). SMARTA's tablet and phone apps are fully functional with no food dispenser, but adding a food dispenser appropriate for the animal subjects in the target study will be necessary for training and testing.

Apparatus cabinet The SMARTA touchscreen tablet is housed in a laser-cut acrylic cabinet. Such a cabinet is inexpensive to produce, yet is durable under frequent abuse from animal subjects (see the first author's GitHub page [<https://github.com/VanceVagell/smarta>] and YouTube channel [https://www.youtube.com/channel/UCBc_N-Bh01Bevsw4aZOJZnA] for detailed instructions). The acrylic cabinet is laser-cut from 1/4-in. black acrylic. Its plans can be cut from three pieces of acrylic: one that is 24 × 24 in. and two that are 24 × 36 in. (see the SMARTA User Manual, [Appendix B](#) in the supplemental materials). From these three pieces of acrylic, nine separate pieces are laser cut in order to assemble the final apparatus cabinet. When fully constructed, the cabinet is a triangular-shaped prism with right triangles on the sides; the base, front, and top flap are rectangles (Fig. 2).

Most of the enclosure is assembled with screws and bolts, and it can easily be constructed in under an hour. The top flap, which is lifted to restock the food rewards, should be attached with black duct tape (used as a flexible hinge). The conveyor belt assembly sits on large metal brackets that are bolted to the front. The tablet is attached to the viewing window from the back using black duct tape; various plastic and metal mounts were explored but none were sufficiently durable. Four rubber feet are attached to the bottom of the enclosure with nuts and bolts.

Software

The SMARTA software was developed to facilitate visual discrimination and choice studies in medium-sized, nonhuman primates. Broadly speaking, it presents a number of visual touch targets and rewards subjects when they make the desired choice. The SMARTA software can be installed as detailed in the SMARTA User Manual (supplementary materials, [Appendix B](#)).

The SMARTA software consists of the phone and tablet applications, written in Java using the Android SDK (software development kit). The source code for this software is open-source and can be accessed online through GitHub (<https://github.com/VanceVagell/smarta>), an online source code repository. The phone application has two main purposes: (1) to act as a remote control for the researcher and (2) to log data sent from the tablet application. The tablet application is primarily used by the subjects while they are being trained or undergoing testing sessions (described below).

From the Android smartphone, the user presses the SMARTA icon to start the application. The initial home screen is the basic data collection screen, with a list of animal subjects and the option to start training or testing (Fig. 3). Users also have the options to calibrate colors (details below), select user accounts and spreadsheets, and manually operate the conveyor belt and food counts (e.g., to recover from a stuck food reward).

SMARTA setup

SMARTA is placed inside the testing enclosure, which in our case consisted of walls made from cement or wire mesh, prior to the subject's arrival. SMARTA is installed against a wall such that the subjects always face away from researchers, to

avoid distraction and/or observer biases (e.g., Iredale, Nevill, & Lutz, 2010; Marsh & Hanlon, 2007; Nickerson, 1998). The researcher stands outside the enclosure and monitors training and initial testing sessions from afar. Food rewards are dispensed manually during training (via the remote-control phone application) in order to reinforce desirable behaviors, and automatically during testing phases.

Capabilities

Calibration mode

Color representation (i.e., the spectra of red, green, and blue light) varies among electronic screens across models and manufacturers, even among those of the same model (Westland & Ripamonti, 2004). It is also important to note that radiance (i.e., the brightness or intensity of color) also varies among devices; light output increases by approximately the square (i.e., “gamma”) of the RGB value. SMARTA thus offers a calibration mode that allows users to specify (and standardize) both the spectrum and intensity functions of each color being presented as a stimulus (Fig. 4). The spectral properties of the stimulus should be measured with a scanning spectral radiometer (e.g., Photo Research Spectrascan 670) that specifies the radiance of the tablet colors in small wavelength steps (in this case, 2 nm) across the visible spectrum. In this way, RGB values can then be entered in SMARTA's calibration mode in order to control and standardize the color values being displayed. Although this is important in any visual discrimination study, precise control of wavelength and intensity output is especially necessary for color vision discrimination tasks (see, e.g., Nunez, Shapley, & Gordon, 2017).

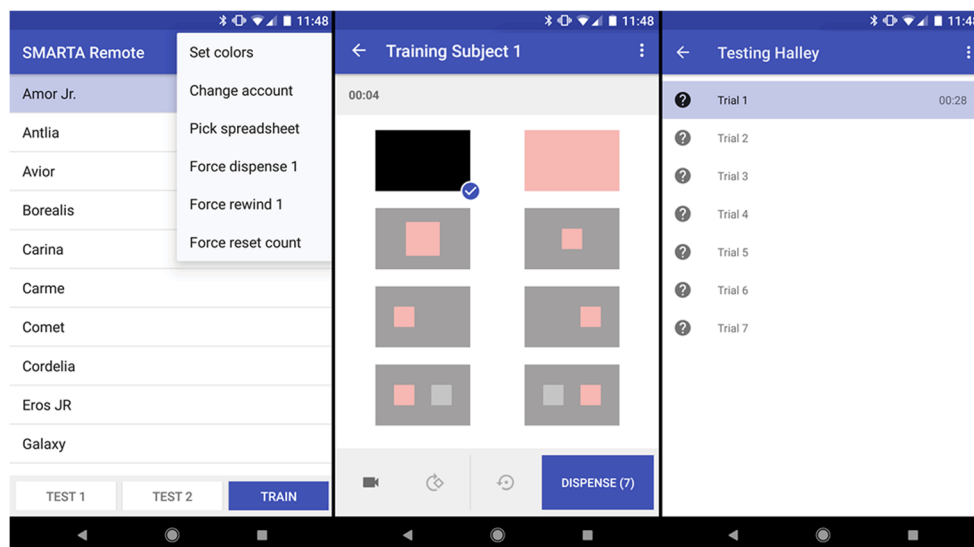


Fig. 3 SMARTA phone application, including the home screen (left), training mode (center), and testing mode (right).

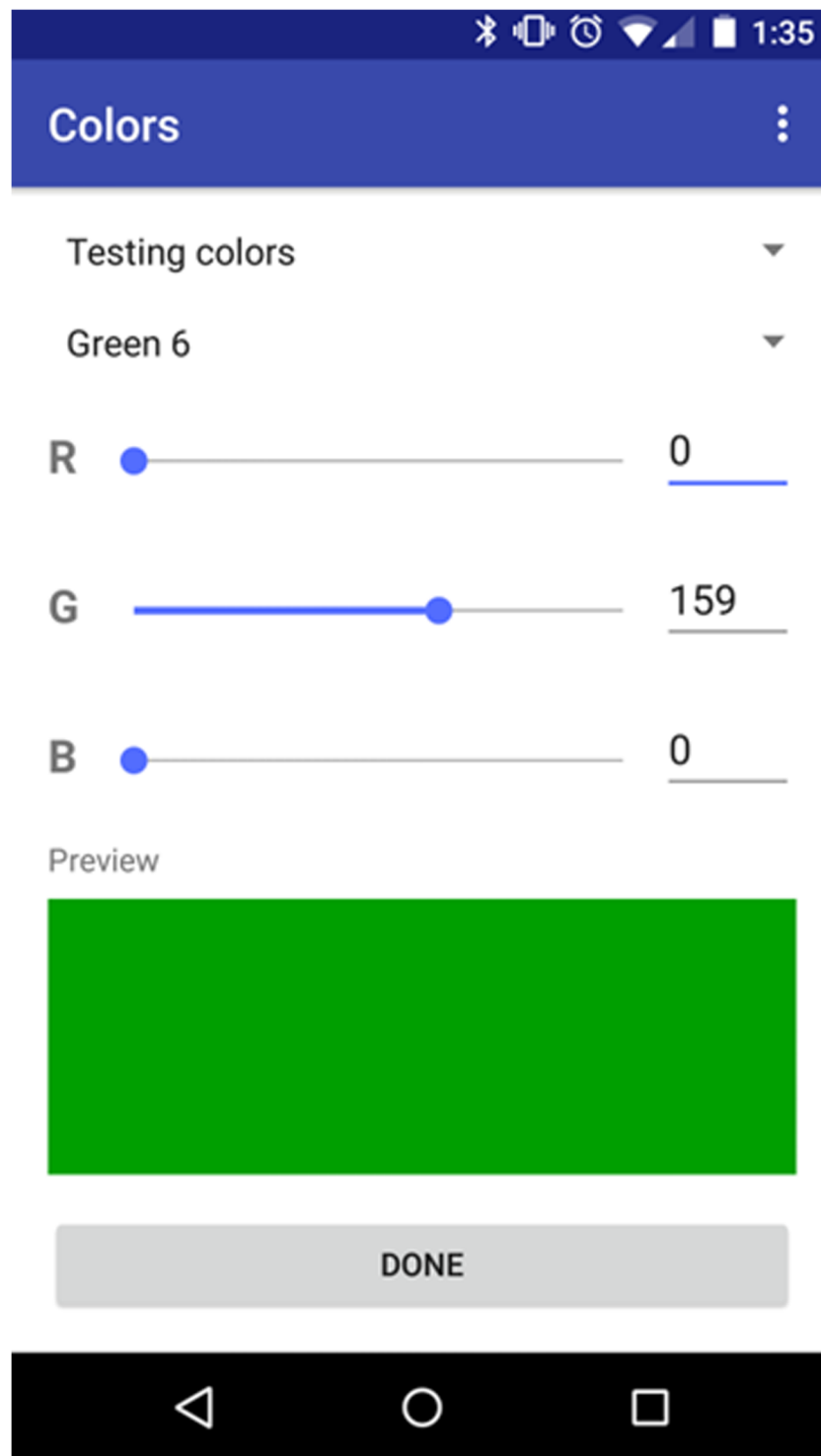


Fig. 4 SMARTA calibration mode for RGB value input.

Video recording

SMARTA automatically records video during testing sessions, beginning at the start of the first trial until the end of the final trial. SMARTA can also manually record videos during training sessions. Videos are recorded from SMARTA's tablet, so it

will record the test subject's behavior from the point of view of SMARTA. Video recordings were not utilized in our case study, but the utilities of this function might include tracking eye movements or assessing the hand preferences of test subjects, among other information. All videos are stored on the tablet for later manual download and analysis.

Training mode

Training mode is a noninteractive mode wherein various visual stimuli can be displayed on the tablet in order to train animal subjects to approach and touch the screen. The screen does not react to the touch of animal subjects; instead, it defers control of all actions to a researcher or trainer via the phone application. Unlike in testing mode, it has no built-in workflow or data capture beyond the subject's name, time, and training duration. After each training trial ends, the data are automatically logged in a Google Sheet.

As in any operant conditioning paradigm, the animals gradually habituate and learn how to use SMARTA through positive reinforcement. SMARTA is programmed to sound a whistle immediately before dispensing food rewards (see the Testing Mode section). There is, however, a brief pause between the time a correct answer is selected and the time it takes for the whistle to sound and the food to be released. In training mode, the whistle can be muted, allowing researchers to manually whistle bridge and dispense food rewards, thereby immediately and precisely reinforcing the desirable behavior.

Testing mode

Testing mode provides an automated workflow with which subjects can interact. In our study, testing mode was designed for two-choice color discrimination tasks (see the Case Study section below). Red was chosen as our standard stimulus, to determine a lemur's ability to discriminate this stimulus on the basis of wavelength, rather than relative brightness, which was made unreliable by varying the brightness of the second stimulus. During Phase 1, SMARTA displays red and gray squares for discrimination (Fig. 5, top). Again, the brightness of the red stimulus remains constant, while the brightness of the gray stimulus varies randomly (seven intensities; see Table 1 for the RGB color values). The positions of red and gray stimuli are randomized during each trial (i.e., left vs. right).

During Phase 2, subjects are tasked to discriminate red from green (Fig. 5, bottom). The brightness of the red stimulus remains constant, while the brightness of the green stimulus varies randomly (seven intensities; see Table 1 for the RGB color values). The positions of the red and green stimuli are randomized during each trial (i.e., left vs. right). When testing mode is activated, video recording automatically begins and trials are executed. After each trial ends, the data are automatically logged in a Google Sheet (Fig. 6).

If the subject selects the correct stimulus during testing mode, SMARTA sounds a preprogrammed whistle tone to bridge the correct response, and the tablet app communicates with the food dispenser via serial connection to release a food reward. There is normally a 5-s pause between trials. However, if SMARTA detects a touch, it will not start the next

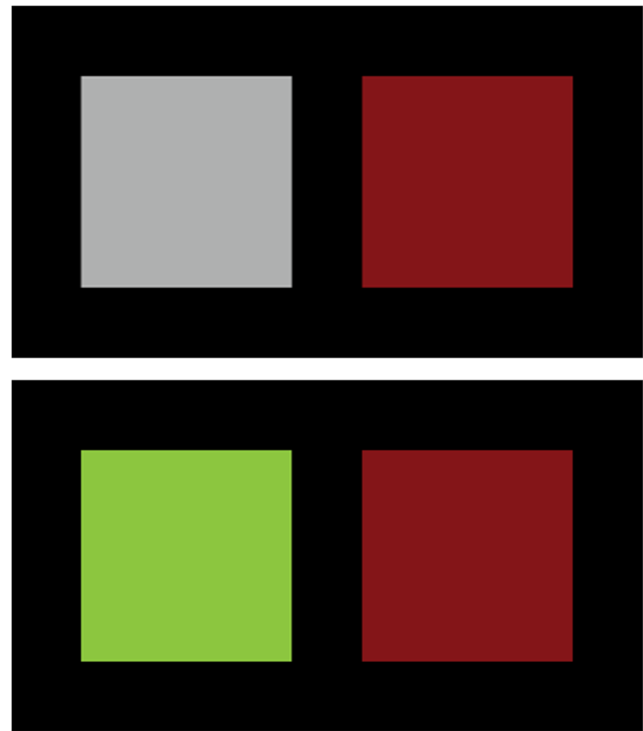


Fig. 5 Example of SMARTA testing mode during Phase 1 (top) and Phase 2 (bottom).

trial until the subject has lifted its hands from the touchscreen. This is to avoid unintentional touches when trials begin.

In this way, SMARTA is a fully automated testing solution when used in testing mode. Once a session is started, it is completely independent of researcher input, and the data are automatically collected and reported to the phone app via Bluetooth for online logging. The researcher can end the session at any time, but otherwise the session will complete without intervention.

Although previous research has included touchscreen discrimination tasks in lemurs (e.g., Joly, Ammersdörfer, Schmidtke, Zimmermann, & Dhenain, 2014; Merritt, MacLean, Crawford, & Brannon, 2011; Merritt, MacLean, Jaffe, & Brannon, 2007), to our knowledge, these apparatuses do not automate data collection and logging in the way that SMARTA does. SMARTA is also programmed to accept correct choices according to specific criteria. SMARTA has custom touch handling; default Android touch handling is disabled. When a touch of any type (i.e., tap, swipe, or long press) is encountered, the x - y coordinates of each finger in the touch are evaluated against the bounds of the two colored squares to check for intersection. If there are intersections, the trial is immediately concluded. The default Android touch handling (e.g., tap handlers on each colored square) was found to be inconsistent, in that subjects would interact with the screen in unexpected ways, such as dragging their hands across it or leaving their hands on it. SMARTA also does not start the next trial until the subject has lifted all fingers

Table 1 RGB color values for red, gray, and green stimuli and their corresponding intensities

TRAINING	R	G	B	Intensity Log ₁₀ (W/m ² ster)	Luminance (cd/m ²)
RED	182	0	0	0	35
GRAY	155	155	155	0.6	7.5
TESTING: PHASE 1	R	G	B	Intensity Log ₁₀ (W/m ² ster)	Luminance (cd/m ²)
RED	182	0	0	0	35
GRAY 1	35	35	35	−0.6	7.5
GRAY 2	45	45	45	−0.4	12
GRAY 3	58	58	58	−0.2	19
GRAY 4	74	74	74	0	30
GRAY 5	94	94	94	0.2	48
GRAY 6	120	120	120	0.4	75
GRAY 7	155	155	155	0.6	120
TESTING: PHASE 2	R	G	B	Intensity Log ₁₀ (W/m ² ster)	Luminance (cd/m ²)
RED	182	0	0	0	35
GREEN 1	0	48	0	−0.6	7.5
GREEN 2	0	62	0	−0.4	12
GREEN 3	0	78	0	−0.2	19
GREEN 4	0	98	0	0	30
GREEN 5	0	125	0	0.2	48
GREEN 6	0	158	0	0.4	75
GREEN 7	0	200	0	0.6	120

off the screen, to prevent errant selections (i.e., before the subject has a chance to see the stimuli). If these criteria are not met, SMARTA will deem a choice incorrect, and thus reduce intra- and interobserver variability, as well as potential researcher bias.

Case study

The SMARTA system was designed to study color vision in ruffed lemurs (*Varecia spp.*) at the Duke Lemur Center (DLC)

in Durham, North Carolina. This colony of ruffed lemurs is known to exhibit a color vision polymorphism, such that some females and all males are red–green color blind (i.e., dichromats), whereas other females have trichromatic color vision (Leonhardt, Tung, Camden, Leal, & Drea, 2008; Rushmore, Leonhardt, & Drea, 2012). At the time of the study, the subjects were housed either in pairs or in groups (up to four individuals) with access to indoor and outdoor spaces that are enclosed by chain-link fencing. Training and testing were restricted to indoor enclosures in order to reduce glare on the touchscreen after animals had been shifted away from other group members to prevent distraction. Because ruffed lemurs are highly food motivated (E. Ehmke, personal communication, July 15, 2014), we delayed morning feedings until sessions were over (up to 1:00 PM) to encourage subjects’ participation in the study. The protocol was approved by Duke University’s Institutional Animal Care and Use Committee (Protocol registry number A252-14-10).

Training

Prior to conducting the experiments, we had used operant conditioning to train subjects to use the SMARTA system. Training was done by chaining and reinforcing the desired behaviors through manual distribution of food rewards. Behaviors were sequentially reinforced (chained) as follows: (a) after approaching SMARTA; (b) after paying attention to the screen; (c) after touching the screen; (d) after touching the red stimulus, when presented with only the red stimulus; and (e) after touching the red stimulus, when it was presented alongside a gray stimulus. Each subject received from two to four training sessions per day on a 4-day-per-week training schedule. Subjects were considered “trained” when they habitually approached SMARTA upon entering the testing area, sat in front of SMARTA facing the screen, waited for the session to begin, and participated in the trials by touching a square when it was displayed on the screen. Subjects that did not meet these criteria after 20 trials were excluded from further training and removed from the study.

	A	B	C	D	E	F	G	H	I	J
1	Subject	Phase	Session start	Trial	Trial duration	Left color	Right color	Timed out	Correct	
2	Halley	Phase 1	2015-07-10 10:30:50 EDT	1	2	R: 124 G: 124 B: 124	R: 191 G: 0 B: 0	FALSE	FALSE	
3	Halley	Phase 1	2015-07-10 10:30:50 EDT	2	3	R: 69 G: 69 B: 69	R: 191 G: 0 B: 0	FALSE	FALSE	
4	Halley	Phase 1	2015-07-10 10:30:50 EDT	3	0	R: 87 G: 87 B: 87	R: 191 G: 0 B: 0	FALSE	FALSE	
5	Halley	Phase 1	2015-07-10 10:30:50 EDT	4	1	R: 191 G: 0 B: 0	R: 77 G: 77 B: 77	FALSE	TRUE	
6	Halley	Phase 1	2015-07-10 10:30:50 EDT	5	1	R: 110 G: 110 B: 110	R: 191 G: 0 B: 0	FALSE	TRUE	
7	Halley	Phase 1	2015-07-10 10:30:50 EDT	6	0	R: 191 G: 0 B: 0	R: 98 G: 98 B: 98	FALSE	TRUE	
8	Halley	Phase 1	2015-07-10 10:30:50 EDT	7	2	R: 139 G: 139 B: 139	R: 191 G: 0 B: 0	FALSE	TRUE	
9										
10										
11										
12										
13										

Fig. 6 Example data sheet for a testing trial (Phase 1 is shown). SMARTA automatically logs the subject information, date and time that training begins, training duration (to the second), RGB color values for left and right squares, whether a trial timed out, and whether a trial choice is correct.

Testing

Once subjects had been trained, they entered two phases of testing. Both phases involved two-choice discrimination tasks. Phase 1 tasked subjects with discriminating between red and gray stimuli. Phase 2 tasked subjects with discriminating between red and green stimuli.

Phase 1 (red vs. gray) was used as a control. Both trichromatic and dichromatic individuals should be able to make discriminations based on achromatic differences of stimuli (Jacobs, 2013; Neitz & Neitz, 2011). Thus, gray was paired with red so that all subjects, regardless of their color vision phenotype, could easily discriminate between the achromatic (gray) and chromatic (red) stimuli. Comparing a constant color in comparison with gray is the standard technique in this field (see, e.g., Abramov & Gordon, 2006).

During Phase 1, the subjects were allowed to enter the SMARTA-equipped research area. The testing session was then initiated remotely using the phone application. Each session consisted of seven trials, with each trial lasting a maximum of 30 s. When a trial ended or when a subject did not make any choice (i.e., the trial timed out), there was a 5-s delay before the next trial began. The maximum time required to run one testing session was therefore 240 s (six 5-s delays and seven 30-s trials). Eight stimuli (seven gray intensities, one red intensity) were randomly paired so as to generate seven different stimulus pairs (colored squares) per session (Table 1). Note that the colors used in this study were calibrated to human spectral sensitivities. However, colors can be calibrated to meet the spectral sensitivities of any trichromatic study subject. For example, our ongoing research with this colony using the SMARTA system has since calibrated colors to meet the known spectral sensitivities of ruffed lemurs (Jacobs & Deegan, 2003; Tan & Li, 1999).

Phase 2 was used to test differences in color perception between trichromatic and dichromatic individuals as the actual testing. Phase 2 operated in the same manner as Phase 1, with the exception that seven green intensities were used instead of gray. Two chromatic stimuli (green and red) were used for Phase 2, in concordance with the hypothesis that trichromats should be able to readily distinguish red from green on the basis of chromatic cues, whereas dichromats should have more difficulty distinguishing between the two stimuli (Gordon & Abramov, 2001; Kaiser & Boynton, 1996). To make brightness cues unreliable, we used seven different gray and green intensities and a constant red intensity during Phases 1 and 2, respectively (Table 1). We maintained red as constant while varying the intensities of gray and green so that intensity cues cannot be consistently used to correctly discriminate red during the testing phase (see Leonhardt et al., 2008). Since the red was maintained at a constant intensity, individuals should associate the red stimulus with a food reward. We therefore expect individuals to select the red

stimulus in order to gain a food reward, unless a study subject had difficulty discriminating red from green (as was expected for dichromats).

During Phase 2, we also introduced an isoluminant pair among the red and green stimuli in order to test the “discrimination based on intensity” hypothesis in dichromats. Isoluminant stimuli have the same intensities and can only be discriminated on the basis of color (Abramov & Gordon, 2006; Teller & Lindsey, 1989). Whereas trichromats should easily discriminate the chromatic difference of an isoluminant pair (Adams, Courage, & Mercer, 1994), dichromats are expected to have more difficulty; they not only perceive the isoluminant pair as having the same color, but also the same intensity, and thus should only discriminate isoluminant pairs at chance level (Clavadetscher, Brown, Ankrum, & Teller, 1988).

Finally, SMARTA was programmed to randomize the stimulus pairs among the trials within each phase of testing. A color selector on the SMARTA phone application specifies the RGB values of all intensities selected for each trial. In Phase 1, SMARTA randomly selected one red intensity and one of seven gray intensities for each trial of testing; in Phase 2, SMARTA randomly selected one red intensity and one of seven green intensities per trial. We tested each animal for 20 sessions in Phase 1 and 20 sessions in Phase 2. On average, each subject required 2 weeks to finish the testing in Phase 1 and Phase 2, averaging three sessions per day on a 4-day-per-week schedule. All sessions were video-recorded for later reference.

Finally, to confirm that SMARTA streamlined data collection and improved the accuracy of results, we asked ten independent observers from our lab to view and score previously recorded testing sessions (Videos 1 and 2) during three independent viewings (i.e., the observers scored six sessions total). We used a combination of Pearson correlations, *t* tests, and maximum likelihood ratios (*G*) to test for differences in our results.

Results

Training The initial training lasted from May to November 2015 and included nine adult ruffed lemurs (five females, four males). Seven of the nine subjects expressed initial interest in the SMARTA system, and five of the nine subjects were trained within five months of initiation ($n = 5$; three females, two males). Subjects that had not been trained after five months were excluded from further participation and were not included in statistical comparisons ($n = 4$; two males, two females). Our final sample included three trichromats and two dichromats.

The subjects required from 42 to 172 training sessions, with an average of 114 sessions required over a four-month period ($n = 5$; $SD = 1.3$, range: 2–5 months), to meet the

acquisition criteria. This was, in part, due to DLC-determined time restrictions when working with animals, as well as logistical constraints due to working with a single SMARTA system (i.e., multiple subjects could not be trained simultaneously). Because training took, on average, 30 min per session per animal, this meant we could only work with a maximum of four animals per day following a 4-day-per-week training schedule. As such, we worked with each subject approximately twice per week.

We found no relationship between total training time and age [$r(4) = .088, p = .888$] or between number of training sessions and age [$r(4) = .048, p = .940$]. Furthermore, we found no significant difference between the sexes in mean total training time [$t(3) = 0.29, p = .605$], nor in mean total number of training sessions [$t(3) = 0.243, p = .824$]. Accordingly, although individuals varied in their levels of interest in training, the total training period for subjects was not influenced by subject sex or age.

Testing During Phase 1 of testing, both the trichromats and dichromats correctly chose red at above-chance levels (more than 50%) for all intensities of gray. Trichromatic and dichromatic individuals did not differ in their overall abilities to discriminate red from gray ($G_2 = 0, p = 1$; Fig. 7). This is in concordance with the expectation that both trichromats and dichromats should be able to discriminate chromatic (red) from achromatic (gray) stimuli (Adams et al., 1994). During Phase 2 of testing, trichromats significantly outperformed dichromats in their ability to discriminate red from green intensities ($G_2 = 78.10, p = .00001$; Fig. 7), choosing red in 73% of trials, versus 45% for dichromats.

When presented with isoluminant pairs (i.e., chromaticity varied, but luminance was the same), the trichromats all performed better than chance, as expected. Of the two dichromats, one performed as expected, selecting the red target approximately half the time (40%); however, the second dichromat (Rees) selected the red target only 9.52% of the time. In other words, at the isoluminant point, Rees consistently chose green over red. This result is unexpected and could be related to our calibration of the colors, which was based on the human isoluminant point. Research is ongoing using color-calibrated stimuli based on lemur absorption curves, as well as a larger sample size of dichromatic and trichromatic individuals.

In addition to demonstrating SMARTA's utility in two-choice color discrimination tasks, our case study demonstrates how the apparatus simplifies and streamlines data collection. Because SMARTA collects data in real time, it took only 156 s (2:36 min) for SMARTA to score and log the results of six sessions; it took human observers an average of 2,186 s (36:26 min) to score the same six sessions ($n = 9; SD = 1,328, \text{range} = 1,200\text{--}5,400 \text{ s}$). This is a more than a 14-fold increase in data collection effort.

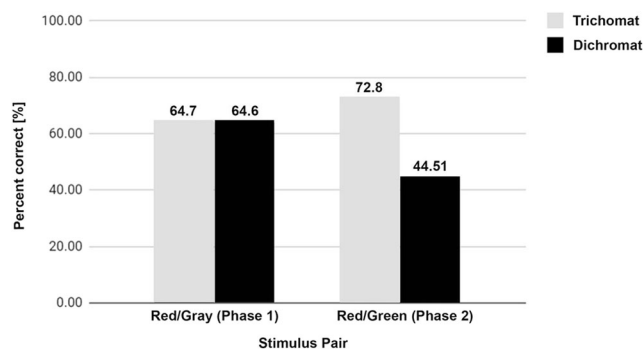


Fig. 7 Percentages of correct choices for trichromats and dichromats during Phase 1 (red and gray stimuli) and Phase 2 (red and green stimuli) of testing. Trichromats and dichromats did not differ significantly when discriminating red and gray stimuli ($G_2 = 0, p = 1$), but trichromats significantly outperformed dichromats when discriminating red and green stimuli ($G_2 = 78.10, p = .00001$).

The SMARTA system also helped improve the accuracy of our study's results. Overall, interobserver agreement was high: 96.4% among the viewers across videos (Video 1: 98.6%; Video 2: 94.3%). However, agreement dropped when comparing the observer scores with SMARTA (overall agreement: 93.6%; 98.6% and 88.6% agreement for Videos 1 and 2, respectively). Although generally still high, human observers did struggle to score some trials, particularly in Video 2 (seven of the ten observers disagreed with SMARTA's assessment for Trial 3). Thus, by automating data collection using prescribed criteria for "correct choices," SMARTA functioned to remove subjective assessments from our study, thereby standardizing the results across test subjects and sessions, as well as across test administrators. This enhances the ability of studies to incorporate volunteers and students with minimal training experience.

SMARTA enhancements and future directions

In the case study presented here, we successfully used SMARTA for a color vision study in captive lemurs; however, SMARTA can be updated by any Android programmer to facilitate other types of visual discrimination studies. For example, SMARTA could potentially be modified to display shapes, photographs, or lexicons. We have filed this suggestion as a feature request on the GitHub project Issue List for future applications.

Moreover, because SMARTA's design is open-source, it can be easily modified to accommodate species and subject specific requirements, including touchscreen dimensions and/or cabinet height. Although raisins and dried cranberries were used as food reward in this case study, any species-specific food reward that would fit inside the dispensing tray can be used. As we noted earlier, other more species-appropriate food dispensing systems could be substituted by making

modifications to SMARTA's electronic components and Arduino software. Finally, although ruffed lemurs typically used their hands to select stimuli, SMARTA can be used by animals that would not normally use their hands, feet, or paws to make a choice. For example, SMARTA could potentially be used by dogs and cats to select stimuli using their noses.

A small modification that has already been made to SMARTA since the data for this article were gathered is to record time measurements in milliseconds, rather than seconds. This change allows for a finer understanding of subject latencies when touching the stimulus.

Summary

SMARTA is a fully integrated subject-mediated automatic remote-testing apparatus with a touchscreen system that can be used for visual discrimination tasks in captive animal studies. The apparatus is inexpensive to assemble and uses freely available applications for the Android operating system (see Github for the source code: <https://github.com/VanceVagell/smarta>). SMARTA collects data automatically, and they can be uploaded online immediately or stored offline for immediate upload once connected. This feature effectively reduces the time spent manually transcribing data. SMARTA's application has been successfully used in a case study of ruffed lemur color vision at Duke Lemur Center.

Author note We thank Duke Lemur Center for making this research and case study possible, especially Erin Ehmke and Meg Dye. We thank our team of research assistants—Stephanie Tepper, Isabel Avery, Abigail Johnson, Christina Del Carpio, David Betancourt, Delaney Davis, Janet Roberts, Cody Crenshaw, Chelsea Southworth, Adrienne Hewitt, Sierra Cleveland, Melaney Mayes, Brandon Mannarino, Miles Todzo, Nicole Crane, Samantha McClendon, Miranda Brauns, and Kristen Moore—for helping with animal training and data collection. We also thank the members of the Hunter College Primate Molecular Ecology Lab (PMEL) for helping us with our validation study: Dominique Raboin, Amanda Puitiza, Aparna Chandrashekar, Jess Knierim, and Christian Gagnon. The case study protocol was approved by Duke IACUC (registry number A252-14-10). This research was funded by the Duke Lemur Center Director's Fund (Publication number 1390).

References

- Abramov, I., & Gordon, J. (2006). Development of color vision in infants. In R. H. Duckman (Ed.), *Visual development, diagnosis, and treatment of the pediatric patient* (pp. 143–170). Philadelphia: Lippincott Williams & Wilkins.
- Adams, R. J., Courage, M. L., & Mercer, M. E. (1994). Systematic measurement of human neonatal color vision. *Vision Research*, *34*, 1691–1701. [https://doi.org/10.1016/0042-6989\(94\)90127-9](https://doi.org/10.1016/0042-6989(94)90127-9)
- Chen, X., & Li, H. (2017). ArControl: An Arduino-based comprehensive behavioral platform with real-time performance. *Frontiers in Behavioral Neuroscience*, *11*, 244. <https://doi.org/10.3389/fnbeh.2017.00244>
- Clavadetscher, J. E., Brown, A. M., Ankrum, C., & Teller, D. Y. (1988). Spectral sensitivity and chromatic discriminations in 3- and 7-week-old human infants. *Journal of the Optical Society of America A*, *5*, 2093. <https://doi.org/10.1364/josaa.5.002093>
- Devarakonda, K., Nguyen, K. P., & Kravitz, A. V. (2016). ROBucket: A low cost operant chamber based on the Arduino microcontroller. *Behavior Research Methods*, *48*, 503–509. <https://doi.org/10.3758/s13428-015-0603-2>
- Emmerton, J., & Delius, J. D. (1980). Wavelength discrimination in the “visible” and ultraviolet spectrum by pigeons. *Journal of Comparative Physiology A*, *141*, 47–52. <https://doi.org/10.1007/BF00611877>
- Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, *41*, 396–404. <https://doi.org/10.3758/BRM.41.2.396>
- Flick, B., Spencer, H., & van der Zwan, R. (2011). Are hand-raised flying-foxes (*Pteropus conspicillatus*) better learners than wild-raised ones in an operant conditioning situation? In B. Law, P. Eby, D. Lunney, & L. Lumsden (Eds.), *The biology and conservation of Australasian bats* (pp. 86–91). Sydney: Royal Zoological Society of New South Wales.
- Gordon, J. & Abramov, I. (2001). Color vision. In: E.B.Goldstein. editor. *The Blackwell Handbook of Perception*. (pp. 92-127). Oxford: Blackwell.
- Hoffman, A. M., Song, J., & Tuttle, E. M. (2007). ELOPTA: A novel microcontroller-based operant device. *Behavior Research Methods*, *39*, 776–782. <https://doi.org/10.3758/BF03192968>
- Iredale, S. K., Nevill, C. H., & Lutz, C. K. (2010). The influence of observer presence on baboon (*Papio* spp.) and rhesus macaque (*Macaca mulatta*) behavior. *Applied Animal Behavior Science*, *122*, 53–57. <https://doi.org/10.1016/j.applanim.2009.11.002>
- Jacobs, G. H. (2013). *Comparative color vision*. St. Louis: Elsevier Science.
- Jacobs, G. H., & Deegan, J. F., II. (2003). Photopigment polymorphism in prosimians and the origins of primate trichromacy. In J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal and defective colour vision* (pp. 14–20). Oxford: Oxford University Press.
- Joly, M., Ammersdorfer, S., Schmidtke, D., & Zimmermann, E., Dhenain, M. (2014) Touchscreen-Based Cognitive Tasks Reveal Age-Related Impairment in a Primate Aging Model, the Grey Mouse Lemur (*Microcebus murinus*). *PLoS ONE* *9* (10):e109393.
- Kaiser, P. K., & Boynton, R. M. (1996). *Human color vision*. Optical Society of America.
- Leonhardt, S. D., Tung, J., Camden, J. B., Leal, M., & Drea, C. M. (2008). Seeing red: Behavioral evidence of trichromatic color vision in strepsirrhine primates. *Behavioral Ecology*, *20*, 1–12. <https://doi.org/10.1093/beheco/arn106>
- Marsh, D. M., & Hanlon, T. J. (2007). Seeing what we want to see: Confirmation bias in animal behavior research. *Ethology*, *113*, 1089–1098. <https://doi.org/10.1111/j.1439-0310.2007.01406.x>
- McSweeney, F. K., Murphy, E. S., Pryor, K., & Ramirez, K. (2014). *The Wiley Blackwell handbook of operant and classical conditioning*. (pp. 455–482). Oxford: Wiley-Blackwell.
- Merritt, D., MacLean, E. L., Crawford, J. C., & Brannon, E. M. (2011). Numerical rule-learning in ring-tailed lemurs (*Lemur catta*). *Frontiers in Psychology*, *2*, 23:1–9. <https://doi.org/10.3389/fpsyg.2011.00023>

- Merritt, D., MacLean, E. L., Jaffe, S., & Brannon, E. M. (2007). A comparative analysis of serial ordering in ring-tailed lemurs (*Lemur catta*). *Journal of Comparative Psychology*, *121*, 363–371.
- Mueller-Paul, J., Wilkinson, A., Aust, U., Steurer, M., Hall, G., & Huber, L. (2014). Touchscreen performance and knowledge transfer in the red-footed tortoise (*Chelonoidis carbonaria*). *Behavioural Processes*, *106*, 187–192. <https://doi.org/10.1016/j.beproc.2014.06.003>
- Müller, C. A., Schmitt, K., Barber, A. L. A., & Huber, L. (2015). Dogs can discriminate emotional expressions of human faces. *Current Biology*, *25*, 601–605. <https://doi.org/10.1016/j.cub.2014.12.055>
- Neitz, J., & Neitz, M. (2011). The genetics of normal and defective color vision. *Vision Research*, *51*, 633–651.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Nunez, V., Shapley, R. M., & Gordon, J. (2017). Nonlinear dynamics of cortical responses to color in the human cVEP. *Journal of Vision*, *17*(11), 9. <https://doi.org/10.1167/17.11.9>
- Oh, J., & Fitch, W. T. (2017). CATOS (Computer Aided Training/Observing System): Automating animal observation and training. *Behavior Research Methods*, *49*, 13–23. <https://doi.org/10.3758/s13428-015-0694-9>
- Pineño, O. (2014). ArduiPod Box: A low-cost and open-source Skinner box using an iPod Touch and an Arduino microcontroller. *Behavior Research Methods*, *46*, 196–205. <https://doi.org/10.3758/s13428-013-0367-5>
- Ribeiro, M. W., Neto, J. F., Morya, E., Brasil, F. L., & de Araújo, M. F. (2018). OBAT: An open-source and low-cost operant box for auditory discriminative tasks. *Behavior Research Methods*, *50*, 816–825. <https://doi.org/10.3758/s13428-017-0906-6>
- Rushmore, J., Leonhardt, S. D., & Drea, C. M. (2012). Sight or scent: Lemur sensory reliance in detecting food quality varies with feeding ecology. *PLoS ONE*, *7*, e41558. <https://doi.org/10.1371/journal.pone.0041558>
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton Century.
- Steurer, M. M., Aust, U., & Huber, L. (2012). The Vienna comparative cognition technology (VCCT): An innovative operant conditioning system for various species and experimental procedures. *Behavior Research Methods*, *44*, 909–918. <https://doi.org/10.3758/s13428-012-0198-9>
- Tan, Y., & Li, W. (1999). Trichromatic vision in prosimians. *Nature*, *402*, 35–36.
- Teller, D. Y., & Lindsey, D. T. (1989). Motion nulls for white vs. isochromatic gratings in infants and adults. *Journal of the Optical Society of America A*, *6*, 1945–1954.
- Westland, S., & Ripamonti, C. (2004). *Computational colour science using MATLAB*. Hoboken: Wiley.